MOLECULAR BIOLOGY OF THE CELL

SECOND EDITION

Bruce Alberts • Dennis Bray
Julian Lewis • Martin Raff • Keith Roberts
James D. Watson

Garland Publishing, Inc.
New York & London

TEXT EDITOR: Miranda Robertson

GARLAND STAFF

Managing Editor: Ruth Adams
Project Editor: Alison Walker
Production Coordinator: Perry Bessas
Designer: Janet Koenig
Copy Editors: Lynne Lackenbach and Shirley Cobert
Editorial Assistant: Māra Abens
Art Coordinator: Charlotte Staub
Indexer: Maija Hinkle

Bruce Alberts received his Ph.D. from Harvard University and is currently Chairman of the Department of Biophysics and Biochemistry at the University of California Medical School in San Francisco. *Dennis Bray* received his Ph.D. from the Massachusetts Institute of Technology and is currently a Senior Scientist in the Medical Research Council Cell Biophysics Unit at King's College London. *Julian Lewis* received his D.Phil. from Oxford University and is currently a Senior Scientist in the Imperial Cancer Research Fund Developmental Biology Unit, Dept. of Zoology, Oxford University. *Martin Raff* received his M.D. degree from McGill University and is currently a Professor in the Biology Department at University College London. *Keith Roberts* received his Ph.D. from Cambridge University and is currently Head of the Department of Cell Biology at the John Innes Institute, Norwich. *James D. Watson* received his Ph.D. from Indiana University and is currently Director of the Cold Spring Harbor Laboratory. He is the author of *Molecular Biology of the Gene* and, with Francis Crick and Maurice Wilkins, won the Nobel Prize in Medicine and Physiology in 1962.

Library of Congress Cataloging-in-Publication Data

Molecular biology of the cell / Bruce Alberts ... [et al.].—2nd ed.
p. cm.
Includes bibliographies and index.
ISBN 0-8240-3695-6.—ISBN 0-8240-3696-4 (pbk.)
1. Cytology. 2. Molecular biology. I. Alberts, Bruce.
[DNLM: 1. Cells. 2. Molecular Biology. QH 581.2 M718]
QH581.2.M64 1989
574.87—dc19
DNLM/DLC
for Library of Congress 88-38275
CIP

Published by Garland Publishing, Inc.
136 Madison Avenue, New York, NY 10016

Printed in the United States of America

15 14 13 12 11 10 9 8 7 6 5 4 3 2

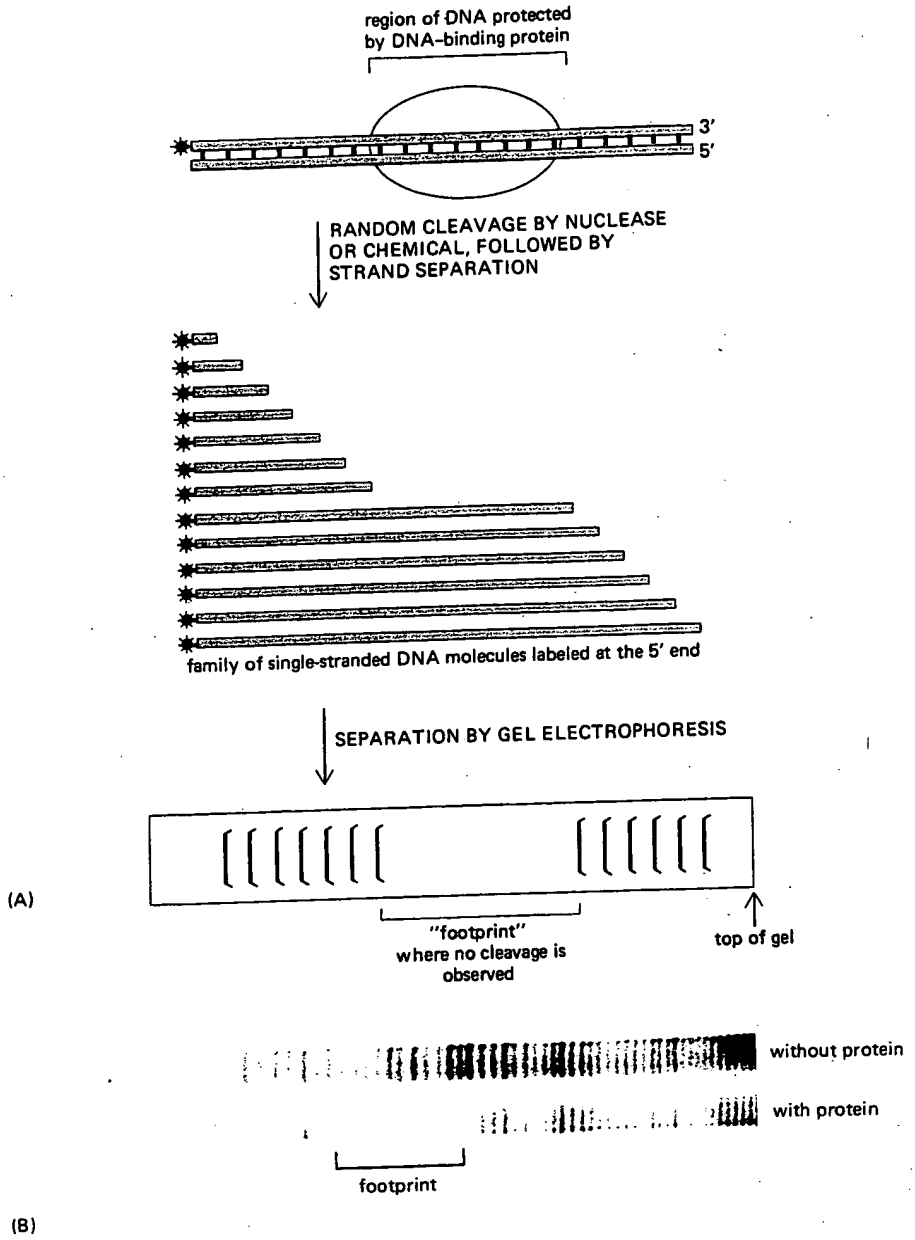


Figure 4–69 The DNA footprinting technique. (A) A protein binds tightly to a specific DNA sequence that is eight nucleotides long, thereby protecting these eight nucleotides from the cleaving agent. If the same reaction were carried out without the DNA-binding protein, a complete ladder of bands would be seen on the gel (not shown). (B) An actual footprint used to determine the binding site for a human protein that stimulates the transcription of specific eucaryotic genes. These results locate the binding site about 60 nucleotides upstream from the start site for RNA synthesis. The cleaving agent was a small, iron-containing organic molecule that normally cuts at every phosphodiester bond with nearly equal frequency. (B, courtesy of Michele Sawadogo and Robert Roeder.)

Nucleic Acid Hybridization Reactions Provide a Sensitive Way of Detecting Specific Nucleotide Sequences[43]

When an aqueous solution of DNA is heated at 100°C or exposed to a very high pH (pH $\geq$ 13), the complementary base pairs that normally hold the two strands of the double helix together are disrupted and the double helix rapidly dissociates into two single strands. This process, called *DNA denaturation*, was for many years thought to be irreversible. However, in 1961 it was discovered that complementary single strands of DNA will readily re-form double helices (a process called **DNA renaturation** or **hybridization**) if they are kept for a prolonged period at 65°C. Similar hybridization reactions will occur between any two single-stranded nucleic acid chains (DNA:DNA, RNA:RNA, or RNA:DNA), provided that they have a complementary nucleotide sequence.

The rate of double-helix formation during hybridization reactions is limited by the rate at which two complementary nucleic acid chains happen to collide, which depends on their concentration in the solution. Hybridization rates can therefore be used to determine the concentration of any desired DNA or RNA sequence in a mixture of other sequences. This assay requires a pure single-

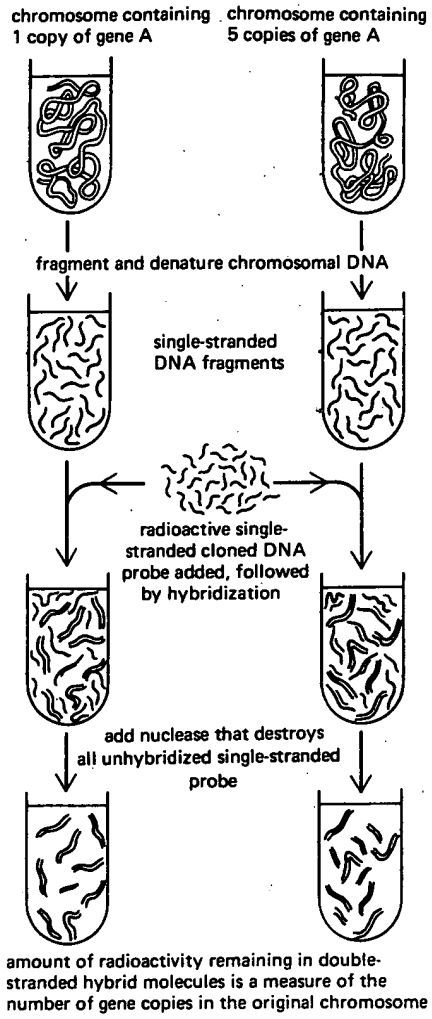


Figure 4–70 Measurement of the number of copies of a specific gene in a sample of DNA by means of DNA hybridization. The radioactive single-stranded DNA fragment used in such experiments is commonly referred to as a *DNA probe;* the chromosomal DNA is not radioactively labeled here.

stranded DNA fragment that is complementary in sequence to the nucleic acid (DNA or RNA) one wishes to detect; the DNA can be obtained by cloning, or if the sequence is short, it can be synthesized by chemical means. In either case the DNA fragment is heavily radiolabeled with ^{32}P (see Figure 4–65) so that its incorporation into double-stranded molecules can be followed during the course of a hybridization reaction. A single-stranded DNA molecule used as an indicator in this way is known as a **DNA probe;** it can be anywhere from 15 to thousands of nucleotides long.

Hybridization reactions using DNA probes are so sensitive and selective that complementary sequences present at a concentration as low as one molecule per cell can be detected (Figure 4–70). It is thus possible to determine how many copies of a particular DNA sequence (contained in the probe) are present in a cell's genome. The same technique can be used to search for related but nonidentical genes; for example, once an interesting gene has been cloned from a mouse or a chicken, part of its sequence can be used as a probe to find the corresponding gene in a human.

Alternatively, DNA probes can be used in hybridization reactions with RNA rather than DNA to find out whether a cell is expressing a given gene. In this case a DNA probe that contains part of the gene's sequence is hybridized with RNA purified from the cell in question to see whether the RNA includes molecules matching the probe DNA and, if so, in what quantities. In somewhat more elaborate procedures the DNA probe is treated with specific nucleases after the hybridization is complete to determine the exact regions of the DNA probe that have paired with cellular RNA molecules. One can thereby determine the start and stop sites for RNA transcription (Figure 4–71); in the same way, one can identify the precise boundaries of the regions that are cut out of the RNA transcripts by *RNA splicing* (the intron sequences—see p. 533).

Large numbers of genes are switched on and off in elaborate patterns as an embryo develops. The hybridization of DNA probes to cellular RNAs allows one to determine whether a particular gene is off or on; moreover, when the expression of a gene changes, one can determine whether the change is due to controls that act on the transcription of DNA, the splicing of the gene's RNA, or the translation of its mature mRNA molecules into protein. Hybridization methods are in such wide use in cell biology today that it is difficult to imagine what it would be like to study gene structure and expression without them.

Northern and Southern Blotting Facilitate Hybridization with Electrophoretically Separated Nucleic Acid Molecules[44]

DNA probes are often used in conjunction with gel electrophoresis to detect the nucleic acid molecules with sequences that are complementary to all or part of the probe. The electrophoresis fractionates the many different RNA or DNA mole-

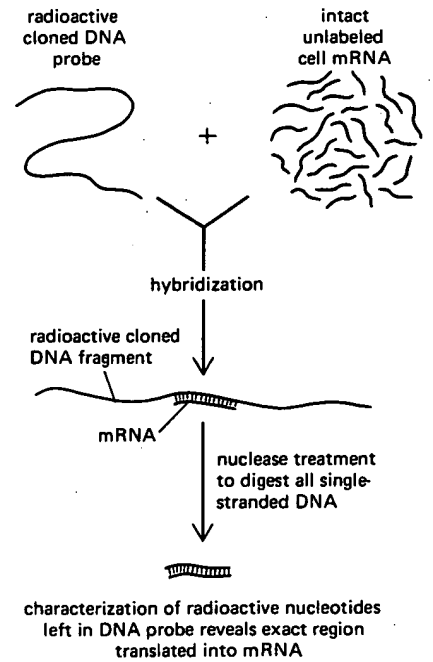


Figure 4–71 The use of nucleic acid hybridization to determine the region of a cloned DNA fragment that is transcribed into mRNA. The method shown requires a nuclease that cuts the DNA chain only where it is not base-paired to a complementary RNA chain. Both the beginning and the end of an RNA molecule can be exactly mapped in this way; in addition, the positions of introns (intervening sequences) in eucaryotic genes are mapped by similar procedures.

cules in a crude mixture according to their size before the hybridization reaction is carried out; if molecules of only one or a few sizes become labeled with the probe, one can be certain that the hybridization was indeed specific. Moreover, the size information obtained can be invaluable in itself. An example will illustrate this point.

Suppose that one wishes to determine the nature of the defect in mutant mice that produces abnormally low amounts of *albumin*, a protein that liver cells normally secrete into the blood in large amounts. First one collects identical samples of liver tissue from defective and normal mice (the latter serving as controls) and disrupts the cells in a strong detergent to inactivate cellular nucleases that might otherwise degrade the nucleic acids. Next one separates the RNA and DNA from all of the other cell components: the proteins present are completely denatured and removed by repeated extractions with phenol—a potent organic solvent that is partly miscible with water; the nucleic acids, which remain in the aqueous phase, are then precipitated with alcohol to separate them from the small molecules of the cell. Then one separates the DNA from the RNA by their different solubilities in alcohols and degrades any contaminating nucleic acid of the unwanted type by treatment with highly specific enzymes—either RNase or DNase.

To analyze albumin-encoding RNAs with a DNA probe for such RNA, a technique called **Northern blotting** is used. First, the intact RNA molecules from defective and control liver cells are fractionated into a series of bands by gel electrophoresis. Then, to make the RNA molecules accessible to DNA probes, a replica of the gel is made by transferring ("blotting") the fractionated RNA molecules onto a sheet of nitrocellulose or nylon paper. The RNA molecules that hybridize to the radioactive DNA probe (because they contain part of the normal albumin gene sequence) are then located by incubating the paper with a solution containing the probe and detecting the hybridized probe by autoradiography (Figure 4–72). Because small nucleic acid molecules move more rapidly through the gel than large ones, the size of each RNA band that binds the probe can be determined by reference to the rates of migration of RNA molecules of known size

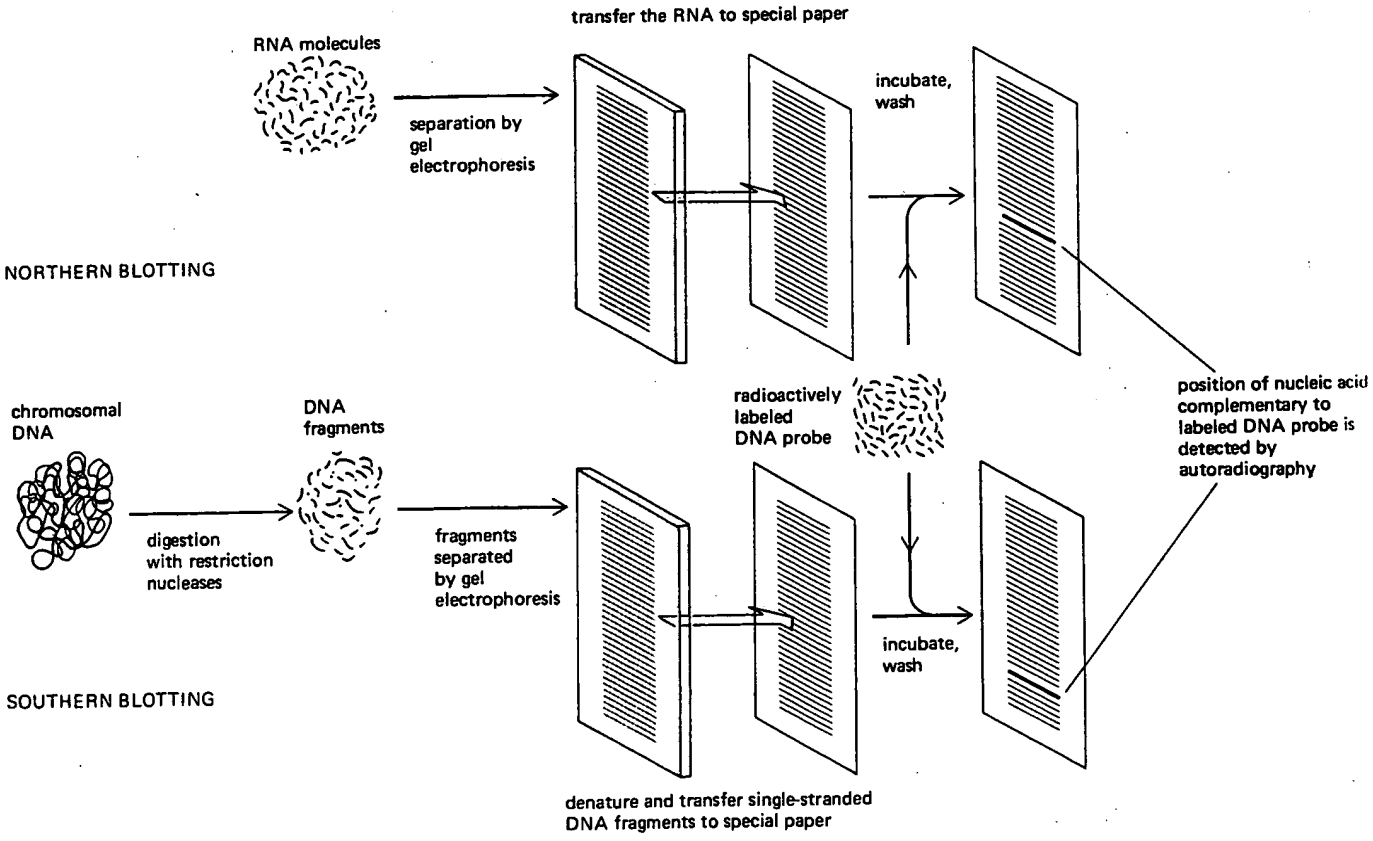


Figure 4–72 Northern and Southern blotting analyses. After the indicated mixture of RNA or DNA molecules is fractionated by electrophoresis through an agarose gel, the many different RNA or DNA molecules present are transferred to nitrocellulose or nylon paper by blotting. The paper sheet is then exposed to a radioactive DNA probe for a prolonged period under hybridization conditions. The sheet is washed thoroughly afterward, so that only those immobilized RNA or DNA molecules that hybridize to the probe become radioactively labeled and show up as bands on autoradiographs of the paper sheet.

make albumin RNA in normal amounts and of normal size; alternatively, normal albumin RNA might be detected in greatly reduced amounts. Another possibility is that the mutant albumin RNA might be abnormally short and therefore move unusually quickly through the gel, in which case the gel blot could be retested with more selective DNA probes to reveal what part of the normal RNA is missing.

To characterize the structure of the albumin gene in the defective mice, an analogous method, called **Southern blotting**, which analyzes DNA rather than RNA, is used. Isolated DNA is first cut into readily separable fragments with restriction nucleases. The fragments are then separated according to their size by gel electrophoresis, and those complementary to the albumin DNA probe are identified by blotting and hybridization, as just described for RNA (see Figure 4–72). By repeating this procedure using different restriction nucleases, a detailed *restriction map* could be constructed for the genome in the region of the albumin gene (see p. 182). From this map one could determine if the albumin gene has been rearranged in the defective animals—for example, by the deletion or the insertion of a short DNA sequence.

Synthetic DNA Molecules Facilitate the Prenatal Diagnosis of Genetic Diseases[45]

At the same time that microbiologists were developing DNA cloning techniques, organic chemists were improving the methods for synthesizing short DNA chains. Today such *DNA oligonucleotides* are routinely produced by machines that can automatically synthesize any sequence up to 80 nucleotides long overnight. This ability to produce DNA molecules of a desired sequence makes it possible to redesign genes at will, an important aspect of genetic engineering, as will be explained in Chapter 5 (p. 266).

Another important use for DNA oligonucleotides is in the prenatal diagnosis of genetic diseases. More than 500 human genetic diseases are attributable to single-gene defects. In most of these the mutation is recessive: that is, it is harmful only when an individual inherits a defective copy of the gene from both parents. One goal of modern medicine is to identify those fetuses that carry two bad copies of the affected gene long before birth so that the mother, if she wishes, can have the pregnancy terminated. For example, in sickle-cell anemia the exact nucleotide change in the mutant gene is known (the sequence GAG is changed to GTG in the DNA strand that codes for the β chain of hemoglobin). For prenatal diagnosis, two DNA oligonucleotides are synthesized—one corresponding to the normal gene sequence in the region of the mutation and the other corresponding to the mutated sequence. By keeping these sequences short (about 20 nucleotides) and selecting a hybridization temperature where only the perfectly matched helix is stable, they can be used as radioactive probes to distinguish between the two forms of the gene. The diagnosis involves isolating DNA from fetal cells collected by amniocentesis and then using the oligonucleotide probes for Southern blotting (see Figure 4–72). A defective fetus can be readily recognized because its DNA will hybridize *only* with the oligonucleotide that is complementary to the mutant DNA sequence. For many genetic abnormalities, the exact nucleotide sequence change is not known. For an increasing number of these, prenatal diagnosis is still possible by using Southern blotting to assay for specific variations in the human genome (called *restriction fragment length polymorphisms*, or *RFLPs*) that are known to be closely linked to the defective gene.

Hybridization at Reduced Stringency Allows Distantly Related Genes to Be Identified[46]

New genes arise during evolution by the duplication and divergence of old genes and by the reutilization of portions of old genes in new combinations (p. 599). For this reason, most genes have a family of close relatives elsewhere in the genome, and some of them are likely to have a related function. Laborious methods are usually required to isolate a DNA clone corresponding to the first member of such

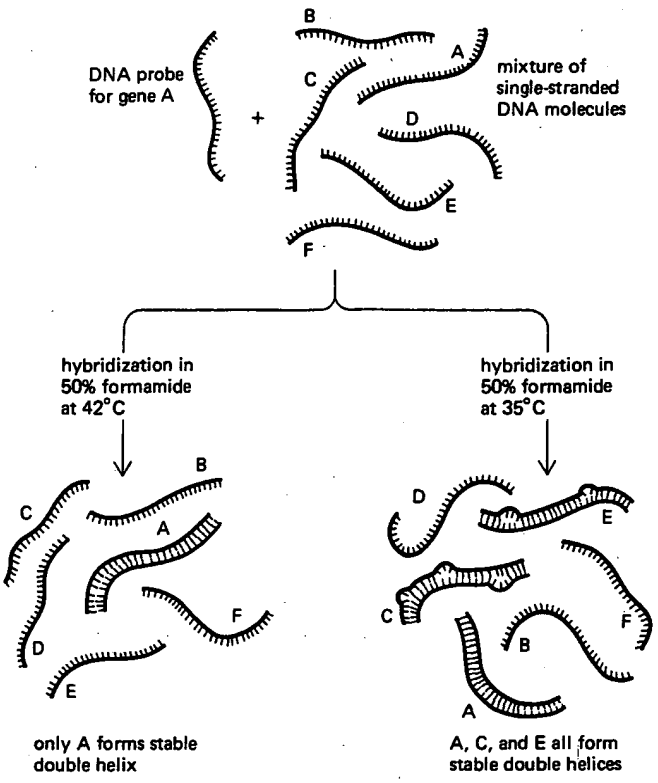


Figure 4–73 Comparison of stringent and reduced-stringency hybridization conditions. In the reaction on the left (stringent conditions), the solution is kept only a few degrees below the temperature at which a perfect DNA helix denatures (its *melting temperature*), so that the imperfect helices that can form under the conditions of reduced stringency on the right are unstable. Only the hybridization conditions on the right can be used to find genes that are nonidentical but related to gene A.

a *gene family* (see p. 262). However, additional genes that produce proteins with related functions can often be isolated relatively easily by using sequences from the first gene as DNA probes. Because the new genes are unlikely to have *identical* sequences, hybridizations with the DNA probe are usually carried out under conditions of "reduced stringency"—defined as conditions that allow even an imperfect match with the probe sequence to form a stable double helix (Figure 4–73).

Although using reduced stringency for hybridization carries the risk of obtaining a false signal from a chance region of short sequence homology in an unrelated DNA sequence, such hybridization represents one of the most powerful uses of recombinant DNA technology. For example, this approach has led to the isolation of a whole family of DNA-binding proteins that function as master regulators of gene expression during embryonic development in *Drosophila* (p. 937). It has also made it possible to isolate members of this same gene family from a variety of other organisms, including humans.

In Situ Hybridization Techniques Locate Specific Nucleic Acid Sequences in Chromosomes and Cells[47]

Nucleic acids, no less than other macromolecules, occupy precise positions in cells and tissues, and a great deal of potential information is lost when these molecules are extracted by homogenization. For this reason, techniques have been developed in which nucleic acid probes are used in much the same way as labeled antibodies to locate specific nucleic acid sequences *in situ*, a procedure called ***in situ* hybridization.** This can now be done both for DNA in chromosomes and for RNA in cells. Highly radioactive nucleic acid probes can be hybridized to chromosomes that have been exposed briefly to a very high pH to disrupt their DNA base pairs. The chromosomal regions that bind the radioactive probe during the hybridization step are visualized by autoradiography. The spatial resolution of this technique can be improved by labeling the DNA probes chemically instead of radioactively. Most commonly, the probes are synthesized with nucleotides that contain a biotin side chain (see p. 177), and the hybridized probes are detected by staining with a network of streptavidin and some type of marker molecule (Figure 4–74).

library all the clones covering the region of the genome that contains the mutant gene of interest. These clones would then be used to make DNA probes to locate the altered gene precisely. Many of the mutant genes that cause genetic diseases in humans will eventually be isolated in this way.

In Vitro Translation Facilitates Identification of the Correct DNA Clone[61]

Despite the power of *in situ* hybridization methods to find specific cDNA or genomic clones from a DNA library, they usually pick out many "false positive" clones. Further ingenuity is required to discriminate between these and authentic positive clones. The task is easiest when the desired clone encodes a protein that has already been characterized by other means. In this case, each candidate cloned DNA fragment is used to purify its complementary mRNA molecules from a mixture of cellular mRNAs by a process called **hybrid selection,** in which an excess of the DNA fragment is separated into single strands and immobilized on a filter that is then used to select complementary mRNA molecules by RNA-DNA hybridization (Figure 5–86). The mRNA purified this way is then allowed to direct protein synthesis in a cell-free system using radioactive amino acids. Finally, the radioactive protein produced is characterized and compared with the expected protein product of the desired clone. A match in such a test is normally a prerequisite for concluding that a cloned DNA fragment encodes the given protein.

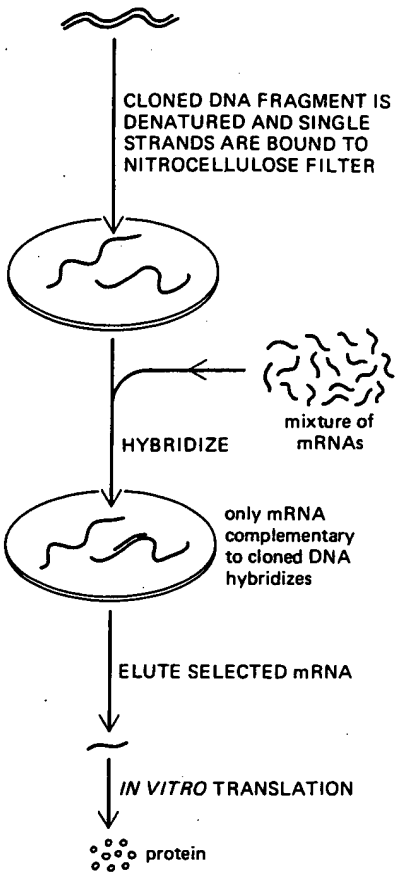


Figure 5–86 The technique of hybrid selection. The purified mRNA molecules are eluted from the filter by subjecting it to conditions that separate the two strands of an RNA-DNA helix.

Expression Vectors Allow cDNA Clones to Be Used to Overproduce Proteins[62]

Very often there is no biochemical information about the protein encoded by a cloned DNA fragment. This is usually the case, for example, when the clone has been identified by subtractive hybridization or by chromosome walking to a mutant gene. In these cases, moreover, the mRNA for the protein in question is often present in such low abundance and in such a limited group of cells that hybrid selection of the complementary mRNA is not feasible. Other methods must then be used to characterize the protein product of the cloned gene. One method is to synthesize a short protein fragment (an *oligopeptide*) corresponding to the deduced amino acid sequence of the protein product of the sequenced cDNA molecule, and then to raise antibodies against the oligopeptide. In many cases the antibodies will recognize the same amino acid sequence when it occurs as part of the natural protein molecule, and these then provide a means to detect, locate, and purify the protein encoded by the original cDNA. In conjunction with cloning by subtractive hybridization, this immunological approach is a powerful way to identify cell type-specific proteins and investigate the development, properties, and functions of each type of cell in a multicellular organism.

The most direct way to characterize the protein encoded by a cloned cDNA, however, is to allow the cDNA itself to direct protein synthesis in a host cell. Plasmids or viruses used for this purpose are called **expression vectors,** and they are constructed so that the cDNA clone is connected directly to a DNA sequence that acts as a strong promoter for DNA transcription. A variety of expression vectors are available, each engineered to function in the type of cell in which the protein is to be made. By means of such *genetic engineering,* bacteria, yeast, or mammalian cells can be induced to make vast quantities of useful proteins—such as human growth hormone, interferon, and viral antigens for vaccines. Bacterial cells with plasmids or viral vectors that have been engineered in this way are especially adept at protein production, and it is common for an engineered gene to produce more than 10% of the total cell protein. Because the production of such a large quantity of a single protein frequently kills the cell, special "inducible" promoters have been designed to allow transcription of the gene to be delayed until a few hours before the cells are collected for protein isolation. Some plasmid expression vectors, for example, contain a bacteriophage lambda promoter that is kept silent by a heat-sensitive gene-repressor protein; the promoter can be turned on at the desired time by increasing the temperature of the bacteria to 42°C, and large amounts of the protein of interest will suddenly be made.

Drosophila chromosomes if the P element transposase enzyme is also present (see p. 256). To make transgenic fruit flies, therefore, the appropriately modified DNA fragment is injected into a very young fruit fly embryo along with a separate plasmid containing the gene encoding the transposase. When this is done, the injected gene often enters the germ line in a single copy as the result of a transposition event (Figure 5–88).

Direct gene replacements have been achieved recently in mice by a laborious, indirect pathway. A DNA fragment containing a desired mutant gene is first transfected into a special line of pluripotent embryo-derived stem cells that grow in cell culture. After a period of cell proliferation, the rare colonies of cells in which general recombination mechanisms have caused a gene replacement to occur are identified by Southern blotting, and individual cells are then implanted into an early mouse embryo. In this new environment the embryo-derived stem cells often proliferate to produce major portions of a normal mouse (see p. 897). Mice that thereby acquire the gene replacement in their germ line are bred to produce both a male and a female animal, each heterozygous for the gene replacement (that is, they have one normal and one mutant copy of the gene). When these two mice are bred, one-fourth of their progeny will be homozygous for the altered gene. Studies of these homozygotes allow the function of an altered gene to be examined in the absence of the corresponding normal gene.

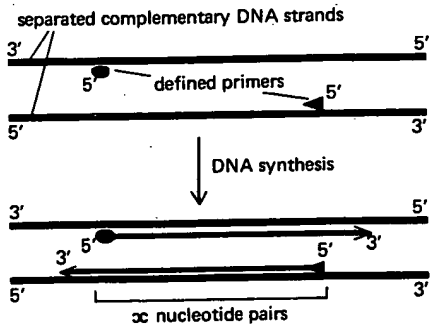
The ability to prepare transgenic mice and *Drosophila* is a powerful new tool for dissecting the function of genes in an intact organism.

Selected DNA Segments Can Be Cloned in a Test Tube by a Polymerase Chain Reaction[65]

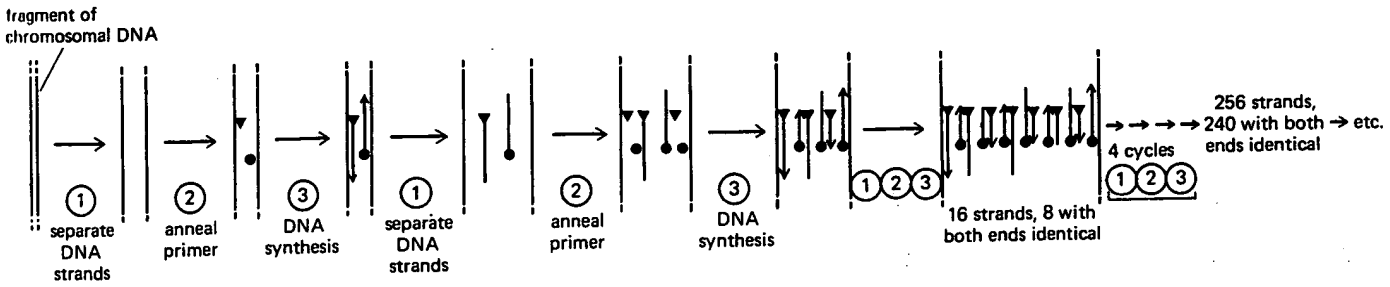
The availability of purified DNA polymerases and chemically synthesized DNA oligonucleotides (see p. 191) has made it possible to clone specific DNA sequences rapidly without the need of a living cell. The technique, called the **polymerase chain reaction (PCR)**, allows the DNA from a selected region of a genome to be amplified by more than a millionfold, provided that at least part of its nucleotide sequence is already known. Portions of the sequence that surround the region to be amplified are used to design two synthetic DNA oligonucleotides, one complementary to each strand of the DNA double helix. These oligonucleotides serve as primers for *in vitro* DNA synthesis, which is catalyzed by a DNA polymerase, and they determine the ends of the final DNA fragment that is obtained (Figure 5–89A).

The principle of the PCR technique is illustrated in Figure 5–89B. Each cycle of the reaction requires a brief heat treatment to separate the two strands of the DNA double helix (step 1). A subsequent cooling of the DNA in the presence of a large excess of the two DNA oligonucleotides allows their specific hybridization to complementary DNA sequences (step 2). The annealed mixture is then incubated with DNA polymerase and the four deoxyribonucleoside triphosphates so that the regions of DNA downstream from the primers are selectively synthesized (step 3). For effective DNA amplification, 20 to 30 cycles of reaction are required. Each cycle doubles the amount of DNA synthesized in the previous cycle. A single cycle requires only about 5 minutes, and an automated procedure permits "cell-free molecular cloning" of a DNA fragment in a few hours, compared to the several days required for standard cloning procedures.

Figure 5–89 The polymerase chain reaction for amplifying specific nucleotide sequences *in vitro*. (A) DNA isolated from cells is heated to separate its complementary strands. These strands are then annealed with an excess of two DNA oligonucleotides (each 15 to 20 nucleotides long) that have been chemically synthesized to match sequences separated by X nucleotides (where X is generally between 50 and 2000). The two oligonucleotides serve as specific primers for *in vitro* DNA synthesis catalyzed by DNA polymerase, which copies the DNA between the sequences corresponding to the two oligonucleotides. (B) After multiple cycles of reaction, a large amount of a single DNA fragment, X nucleotides long, is obtained, provided that the original DNA sample contains the DNA sequence that was anticipated when the two oligonucleotides were designed.



(A)



(B)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.